



PvuI (7)
SgfI (6)
1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGTGCCTA
101 GAGAAGTGGCGCGGGTAAACTGGAAAGTGTGTCGTGACTGGTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

HindIII (245)
Psp1406I (203) **PvuII (239)** **Bsu36I (291)**
201 GTGAACGTTCTTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACCGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCGGTTGAGTCGCGTTTCTGCCGCTCCCGCCTGTGGTGCCTCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

NgoMIV (441)
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCTAACTCTACGCTTTTGTTCGTTT

NcoI (560)
BstEII (555)
KasI (535) **AgeI (552)**
501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGGTACCATTGGCGCCTTCTCCGAGATGGGTGTAATGCCTGAGAT
1 M A A F S E M G V M P E I

EcoRV (640)
601 TGCACAAGCTGTGGAAGAGATGGATTGGCTCCTCCAACCTGATATCCAGGCTGAATCTATCCCATTGATCTTAGGAGGAGGTGATGTACTTATGGCA
13 A Q A V E E M D W L L P T D I Q A E S I P L I L G G G D V L M A A
701 GAAACAGGAAGTGGCAAACTGGTGCTTTAGTATTCCAGTTATCCAGATAGTTTATGAAACTCTGAAAGACCAACAGGAAGGCAAAAAGGAAAAACAA
47 E T G S G K T G A F S I P V I Q I V Y E T L K D Q Q E G K K G K T
801 CAATTAAGCTGGTGCTTCACTGCTGAACAAATGGCAGATGAACCCATATGACAGGATGTGCTTTTGAATTTGGTCAAGTGGTCTTTGTTGTCAAAG
80 T I K T G A S V L N K W Q M N P Y D R G S A F A I G S D G L C C Q S

AseI (942) **BspHI (978)**
901 CAGAGAAGTAAAGGAATGGCATGGGTGTAGAGCTACTAAAGGATTAATGAAAGGAAACACTACTATGAAGTATCCTGTGCATGACCAAGGGTTATGCAGG
113 R E V K E W H G C R A T K G L M K G K H Y Y E V S C H D Q G L C R

StuI (1019) **AvrII (1032)**
1001 GTCGGGTGGTCTACCATGCAGGCTCTTTGGACCTAGGTAAGTACTGACAAGTTGGATTTGGCTTTGGTGGAAACAGGAAAGAAATCCATAACAAACAAATTTG
147 V G W S T M Q A S L D L G T D K F G F G F G G T G K K S H N K Q F

EcoRI (1114) **NsiI (1123)** **BspLU11I (1164)** **BglIII (1191)**
1101 ATAATTATGGAGAGGAATTCACATGTCATGATACCATGGATGTTACCTGGATATAGATAAGGGACATGTCAAGTTCTCCAAAAATGGAAAAGATCTTGG
180 D N Y G E E F T M H D T I G C Y L D I D K G H V K F S K N G K D L G
1201 TCTGGCATTGAAATACCACACATATGAAAAACCAAGCCTCTTCTCGCTGTGTTTTGAAGAATGCTGAACTGAAATTTAACTTCGGTGAAGAGGAA
213 L A F E I P P H M K N Q A L F P A C V L K N A E L K F N F G E E E
1301 TTTAAGTTTCCACAAAAGATGGCTTTGTTGCTCTTCCAAGGCACCGGATGGTTACATTGTCAAATCACAGCACTCAGGTAATGCACAGGTGACACAAA
247 F K F P P K D G F V A L S K A P D G Y I V K S Q H S G N A Q V T Q

XmaI (1445)
1401 CAAAGTTTCTCCCAATGCTCCGAAAGCTCTCATTGTTGAACCTTCCCGGGAGTTAGCTGAACAACTTTGAACAACATCAAGCAGTTTAAAGAAATACAT
280 T K F L P N A P K A L I V E P S R E L A E Q T L N N I K Q F K K Y I

XmaI (1547)
1501 TGATAATCCTAAATTAAGGGAGCTTCTGATAATTGGAGGTGTTGCAGCCGGGATCAGCTCTCTGTTTTGAAAATGGAGTAGATATAGTTGTGGGTA
313 D N P K L R E L L I I G G V A A R D Q L S V L E N G V D I V V G T
1601 CCGGGAAGACTAGATGACTTGGTGTCAACTGAAAAGCTGAACTTATCTCAAGTTAGATTCTGGTCTGGATGAAGCTGATGGCCTTCTTCTCAAGGTT
347 P G R L D D L V S T G K L N L S Q V R F L V L D E A D G L L S Q G

BstEII (1740)
Bsu36I (1735)
1701 ATTCTGATTTTATAAATAGGATGCACAATCAGATTCCTCAGGTTACCTCTGATGGAAAAAGACTTCAGGTGATTGTTTGTCTGCCACTTTGCATTCTTT
380 Y S D F I N R M H N Q I P Q V T S D G K R L Q V I V C S A T L H S F

NsiI (1828) **BsrGI (1881)**
1801 CGATGTAAGAACTGTCCGAGAAGATAATGCTTTTCTACATGGGTTGACTTAAAAGGAGAAGACTCTGTTCCAGATACTGTACACCATGTTGTTGTC
413 D V K K L S E K I M H F P T W V D L K G E D S V P D T V H H V V V

BsrGI (1965)
1901 CCAGTAAATCCAAAAGTACAGACTCTGGAAAGGCTTGGAAAGAGCCACATTAGAAGTATGATGTACATGCAAAAAGATAACACAAGACCTGGTGCTA
447 P V N P K T D R L W E R L G K S H I R T D D V H A K D N T R P G A

BstXI (2005)
2001 ATAGTCCAGAGATGGTCTGAAGCTATTAATCCTGAAAGGGAGTATGCTGTCCGGCAATCAAGGAACATAAGATGGATCAAGCAATTATCTTCTG
480 N S P E M W S E A I K I L K G E Y A V R A I K E H K M D Q A I I F C

ScaI (2132) **EcoO109I (2154)**
2101 TAGAACAAAATTGACTGTGATAACTGGAGCAGTACTTTATACAACAAGGAGGAGGACCTGATAAAAAAGGACACCAGTTCTCATGTGTTTGTCTTCAT
513 R T K I D C D N L E Q Y F I Q Q G G G P D K K G H Q F S C V C L H

BspHI (2214) **EcoRV (2299)**
2201 GGTGACAGAAAGCCTCATGAGAGAAAGCAAACTTGGAAAGATTTAAGAAAGGAGATGTAAGATTCTGATTTGCACAGATGTAGCTGCTAGAGGAATTG
547 G D R K P H E R K Q N L E R F K K G D V R F L I C T D V A A R G I

SnaBI (2356) **SapI (2381)**
2301 ATATCCACGGTGTCTTATGTTATAAATGTCACCTCTGCCGATGAAAAGCAAACTACGTACATCGAATTGGCAGAGTAGGAAGAGCTGAAAGGATGG
580 D I H G V P Y V I N V T L P D E K Q N Y V H R I G R V G R A E R M G

Acc65I (2441)
2401 TCTGGCAATTTCCCTGGTGGCAACAGAAAAAGAAAAGTTGGTACCATGTATGTAGCAGCCGTGAAAAGGGTGTATAACACAAGACTCAAGGAAGAT
613 L A I S L V A T E K E K V W Y H V C S S R G K G C Y N T R L K E D

2501 GGAGGCTGTACCATATGGTACAACGAGATGCAGTTACTATCTGAGATAGAAGAACACCTGAACTGTACCATTTCTCAGGTTGAGCCGGATATAAAGGTAC
 647▶ G G C T I W Y N E M Q L L S E I E E H L N C T I S Q V E P D I K V
 2601 CAGTGGATGAATTTGATGGGAAAGTTACCTACGGTCAGAAAAGGGCTGCTGGTGGGAAAGCTATAAAGGCCATGTGGATATTTTGGCACCTACTGTTCA
 680▶ P V D E F D G K V T Y G Q K R A A G G G S Y K G H V D I L A P T V Q

MscI (2791)

PvuII (2764)

NheI (2785)

2701 AGAGTTGGCTGCCCTTGAAGAGGAGCGCAGACATCTTTCCTGCATCTTGGCTACCTTCTAACAGCTGTTTCAGAACCTTCTGAGCTAGCTGGCCAGAC
 713▶ E L A A L E K E A Q T S F L H L G Y L P N Q L F R T
 2801 ATGATAAGATACATTGATGAGTTTGGACAAACCACAAC TAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAA

HpaI (2923)

2901 CCATTATAAGCTGCAATAAAACAAGTTAAACAACAACAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAA

EcoRI (3019)

3001 CCTCTACAAATGTGGTATGGAATTTCAAATACAGCATAGCAAACCTTAACTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGG

3101 CATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTTAAAGATATAGTGATTTTTCCCAAGTTTGAAGT

SapI (3201)

SspI (3258)

SwaI (3272)

3201 AGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCCACATTCCCTTTTTAGTAAAAATTCAGAAAAATTTAAATACATCATTGCAATGAAAATA

EcoO109I (3333)

3301 AATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAGGAACCTTTAATAGAAT

3401 TGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCATCTCAAT

141

• N R T Y K L P I L E E I T T K V L K G N M E I

SaeI (3533)

BstXI (3562)

3501 GAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGG
 117▶ L V F C D P A Y D S I L E R C M G C P S V V R I S R D V E D S Y P

StuI (3697)

3601 TGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCAGCACAGACAGTACCCTGCCAATGTAGG
 83▶ H R V A V I T D F D K Q G N S V A S G I A I A E A C V T V R G I Y A

3701 CCTCAATGTGGACAGCAGAGATGATCTCCCAGTCTTGGTCTGATGGCCGCCCGACATGGTCTTGTGTCCTCATAGAGCATGGTATCTTCTCAGT
 50▶ E I H V A S I I E G T K T R I A A G V H H K N D E Y L M T I K E T

XmnI (3839)

3801 GGCGACCTCCACCAGTCCAGATCTGCTGAGAGATGTTGAAGTCTTTCATGGTGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGC
 17▶ A V E V L E L D Q Q S I N F T K M

AseI (3905)

SaeI (3962)

3901 CGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTCACTAAACGAGCTCTGCTTATATAGACCTCCCACCGTACACGCCTA

SpeI (4060)

4001 CCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAAAACAACTCCCATTTGACGTCAATGGGGTGGGA

SnaBI (4188)

4101 GACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTA

4201 CTGCCAAGTAGGAAAGTCCATAAGGTCATGTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGA

4301 TACACTTGATGTACTGCCAAGTGGCAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCA

PacI (4479)

PstI (4472)

SdaI (4471)

BspLU11I (4489)

4401 TTATTGACGTCAATGGGGCGGGGCTGTTGGGCGGTGAGCCAGGCGGGCCATTTACCGTAAAGTTATGTAACGCCTGCAGGTTAAATAAGAACATGTGAGCA

4501 AAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCAGAAAAATCGACGCTCAA

4601 GTCAGAGGTGGCAAACCCGACAGGACTATAAAGATACAGGCGTTTTCCCTGGAAGCTCCCTGTCGCTCTCCTGTTCCGACCCTGCCGCTTACCGG

4701 ATACCTGTCCGCTTTCTCCCTTCCGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGATAGTATCTCAGTTCCGTTGATAGTCTGTTCCGCTCAAGCTGGGC

ApaLI (4803)

4801 TGTGTGCAGAACCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGTAGTCCAACCCGTTAAGACACGACTTATCGCCACTGGCAG

4901 CAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATT

5001 TGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGTTTTTTGTT

5101 TGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAACTCACGTTAAG

EagI (5239)

PacI (5219)

SwaI (5228)

NotI (5238)

5201 GGATTTTGGTCATGGCTAGTTAATTAACATTTAAATC AGCGGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCG

5301 TAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAAATAGGCTGTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATC

5401 GAA